Development of an Open-Source Integrated Testing Strategy for Skin Sensitization Potency

J Pirone[1], M Smith[1], N Kleinstreuer[2], T Burns[2], J Strickland[2], Y Dancik[3], R Morris[1], W Casey[4], J Jaworska[3], D Allen[2]

[1]SSS, Inc., Durham, NC, USA; [2]ILS/NICEATM, RTP, NC, USA; [3]P&G NV, Strombeek-Bever, Belgium; [4]NIH/NIEHS/DNTP/NICEATM, RTP, NC, USA

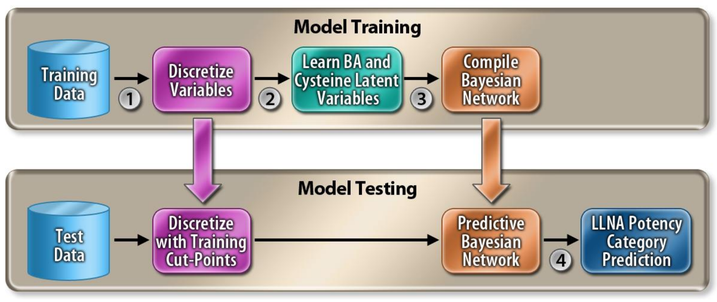
Objective

- To reduce and potentially eliminate animal use for skin sensitization testing, potency results from the LLNA were used as the target endpoint to develop an integrated testing strategy (ITS) using a Bayesian network (BN) (Jaworska et al. 2011, 2013).
- The BN ITS:
  - Combines relevant *in silico* and *in vitro* data to make probabilistic predictions of skin sensitization potency category
  - Aligns with the adverse outcome pathway (AOP) for substances that initiate the skin sensitization process by crossing the skin barrier and covalently binding to skin proteins (OECD 2012)
- The objective of this project was to develop an open-source BN ITS. Previous versions of the versions of the BN ITS (Jaworska et al. 2011, 2013) were developed with commercial software.

Methods

- We developed an open-source (OS) version of the most recent BN ITS (ITS-2) using the free statistical programming language *R* (R v3.0.1, GNU Public License v3) (R Development Core Team 2008) (Pirone et al. 2014).
- Refinements to ITS-2 in the OS version include data corrections and modifications to physicochemical parameters.
- The refined model, OS ITS-2 lipid, uses only the lipid pathway for determining skin bioavailability (http://www.cdc.gov/niosh/topics/skin/finiteSkinPermCalc.html) and is posted on the NTP website at http://ntp.niehs.nih.gov/go/its.
- The process for building and testing OS ITS-2 lipid using high quality *R* packages is shown in **Figure 1**.

Figure 1. Process for Building the OS ITS-2 Lipid Model



Abbreviations: BA = Bioavailability
Variables are defined in **Table 1** and the structure of the network is shown in **Figure 2.**

- Step 1 used the **discretization** package (Kim 2012), which contains implementations of several algorithms for supervised discretization.
- In Step 2, the Bioavailability and Cysteine latent variables were learned using tools from the **poLCA** package (Linzer and Lewis 2011).
- In Step 3, **gRbase** (Dethlefsen and Højsgaard 2005) and **gRain** (Højsgaard, 2012) supply the functions for constructing, parameterizing and performing inference on Bayesian networks.

Database

- The *in vitro* and *in silico* data variables relevant to skin sensitization used to train the model are shown in **Table 1**. The structure of the OS ITS-2 lipid model is shown in **Figure 2**.
- The BN was trained to a dataset consisting of 124 substances with LLNA potency categories distributed as 36 nonsensitizers, 28 weak sensitizers, 35 moderate sensitizers, and 25 strong or extreme sensitizers.
- The LLNA potency predictions of the model were tested using 21 substances in an external test set: 6 nonsensitizers, 5 weak sensitizers, 5 moderate sensitizers, and 5 strong or extreme sensitizers.

Table 1. Variables for the Open-Source ITS-2 Lipid Model

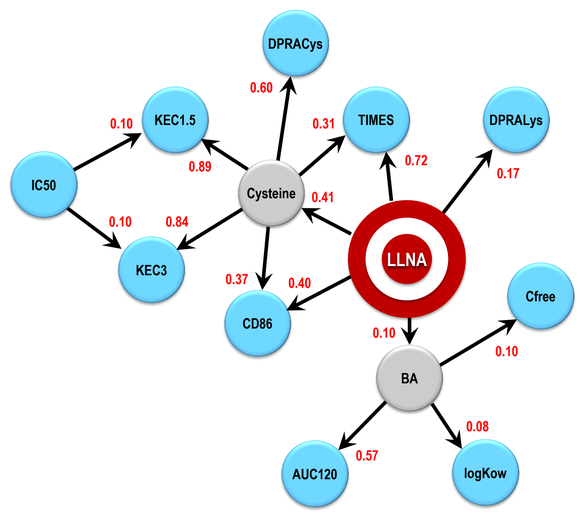
Variable	Description	Measurement	Abbreviation in Figure 2
LLNA	Potency classification in four categories	Nonsensitizer; or weak, moderate, or strong/extreme sensitizer	LLNA
U937 Activation Test	*In vitro* test that uses the human myeloid cell line U937	EC150 (µM) for CD86 cell surface marker expression	CD86
Direct Peptide Reactivity Assay	*In chemico* method that measures peptide remaining after the test substance binds to two model heptapeptides	1) Percent cysteine peptide remaining 2) Percent lysine peptide remaining	1) DPRACys 2) DPRALys
KeratinoSens™ Assay	*In vitro* test that detects electrophiles using the Nrf2 electrophile-sensing pathway in the HaCaT (immortalized keratinocyte) cell line	1) Average concentration that produces 1.5-fold enhanced activity (µM) 2) Average concentration yielding 3-fold enhanced activity (µM) 3) Concentration producing 50% cytotoxicity (µM)	1) KEC1.5 2) KEC3 3) IC50
Physicochemical Property	Octanol–water partition coefficient	Log K_{ow}	logK_{ow}
Bioavailability	Concentration of chemical reaching the mid-epidermal layer of skin calculated using a transdermal transport model (Kasting et al. 2008).	1) Free test substance concentration in mid-epidermis multiplied by thickness of viable epidermis (0.01 cm) (% applied dose) 2) Area under the flux curve at 120 h (% applied dose)	1) Cfree 2) AUC120
TIMES-M	*In silico* categorical prediction of skin sensitization potency using TIMES (Tissue Metabolism Simulator) software (V.2.25.7), an expert system that makes predictions based on knowledge about the parent compound and potential skin metabolites (Dimitrov et al. 2005).	Three categories: nonsensitizer, weak sensitizer, and moderate/strong/extreme sensitizer	TIMES

Abbreviations: EC150 = effective concentration that produces a 1.5-fold increase in the CD86 cell surface marker expression, the threshold for a positive response in the U397 activation test; EC3 = effective concentration that produces a stimulation index of 3, the threshold for a positive response in the LLNA; LLNA = murine local lymph node assay.

References

A reference list for this poster is available at http://ntp.niehs.nih.gov/iccvam/meetings/9wc/pirone-its-refs.pdf

Figure 2. Structure of the OS ITS-2 Lipid



The arrows show the conditional dependencies of the variables that impact murine local lymph node assay (LLNA) potency, which is the target variable. The remaining variables are manifest variables. Bioavailability (BA) and Cysteine are latent variables for bioavailability and cysteine binding, respectively. Mutual information values are shown in red type. The abbreviations for all variables are listed in **Table 1**.

Results

- The LLNA potency category predictions of the OS ITS-2 lipid model using *R* are shown in **Tables 2** and **3** for the training sets and test sets, respectively. The bold red numbers in the tables show the results of the commercial BN ITS-2 lipid model in cases where there is a difference between results it produced and those produced by the OS ITS-2 lipid model.
  - For the training set, the accuracy of LLNA potency category predictions was greater for the OS ITS-2 lipid model: 78% (97/124) vs. 76% (94/124) for the commercial BN ITS-2 lipid model.
    - Using the OS ITS-2 lipid model, 15 substances (12%) were overclassified (predicted category was more severe than observed in the LLNA) and 12 substances (10%) were underclassified (predicted category was less severe than observed in the LLNA).
    - Using the commercial BN ITS-2 lipid model, 21 substances (17%) were overclassified and 9 substances (7%) were underclassified.
  - For the test set, the accuracy of potency category predictions was identical for the OS ITS-2 lipid model: 86% (18/21) vs. 86% (18/21) for the commercial BN ITS-2 lipid model.
    - Using the OS ITS-2 lipid model, no substances were overclassified and 3 substances (14%) were underclassified.
    - Using the commercial BN ITS-2 lipid model, 1 substance (18%) was overclassified and 2 substances (10%) were underclassified.

Table 2. Confusion Matrix for LLNA Potency Category Predictions on the Training Set

Predicted Potency Category[a]	Observed Potency Category[a]			
	Nonsensitizer (36)	Weak Sensitizer (28)	Moderate Sensitizer (35)	Strong/Extreme Sensitizer (25)
Nonsensitizer (36) (32)	31 29	2 1	1	2 1
Weak Sensitizer (27) (26)	3	22 21	2	0
Moderate Sensitizer (35)	1 3	3 4	26 24	5 4
Strong/Extreme Sensitizer (26) (31)	1	1 2	6 8	18 20

Abbreviations: LLNA = murine local lymph node assay.

[a] The numbers in parentheses show the total number of chemicals predicted or observed in each category. Numbers in **bold red** show the different values yielded by the BN ITS-2 lipid model developed using commercial software (Jaworska et al. 2013).

Table 3. Confusion Matrix for LLNA Potency Category Predictions on the Test Set

Predicted Potency Category[a]	Observed Potency Category[a]			
	Nonsensitizer (6)	Weak Sensitizer (5)	Moderate Sensitizer (5)	Strong/Extreme Sensitizer (5)
Nonsensitizer (7)	6	1	0	0
Weak Sensitizer (5) (4)	0	4	1 0	0
Moderate Sensitizer (5)	0	0	4	1
Strong/Extreme Sensitizer (4) (5)	0	0	0 1	4

Abbreviations: LLNA = murine local lymph node assay.

[a] The numbers in parentheses show the total number of chemicals predicted or observed in each category. Numbers in **bold red** show the different values yielded by the BN ITS-2 lipid model developed using commercial software (Jaworska et al. 2013).

Conclusions

- The OS ITS-2 lipid model for skin sensitization potency adequately reproduces the BN ITS-2 lipid model developed using commercial software.
- The open-source model
  - Increases the availability and transparency of the ITS
  - Represents a major step in allowing the ITS to be reproduced and tested, which are essential for implementation in a regulatory framework
- OS ITS-2 lipid is available to the public for testing at http://ntp.niehs.nih.gov/go/its.
- Future work will
  - Substitute the human cell line activation test for the U937 assay
  - Evaluate open source replacements for the TIMES-M *in silico* predictions and open sources for physicochemical properties needed for the bioavailability calculations
  - Add additional substances to the trained model as data are collected

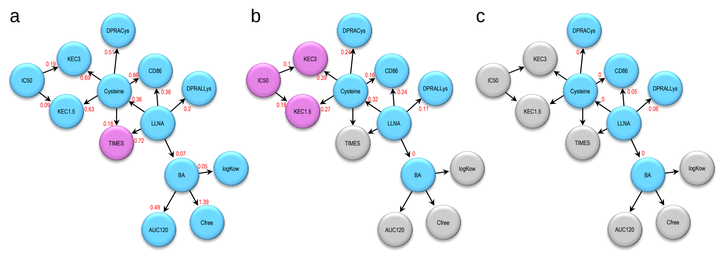
Case Studies

- Chlorobenzene and 2-mercaptobenzothiazole are two case studies that illustrate how the OS ITS-2 lipid model can use existing information to determine the *in vitro* or *in silico* tests that would be most effective for determining the potency classification.
- Value of information (VoI) from all possible sources determines which variable provides the most information about LLNA potency. VoI was assessed by calculating the mutual information between variables, which determines the uncertainty in one variable that is reduced by knowing the results from another variable.

Case Study 1. Chlorobenzene

- Chlorobenzene is a nonsensitizer (ICCVAM 2009).
- Testing Strategy
  - **Figure 3** shows how the mutual information values of the variables change as information is added. The accompanying tables show how the probability for each potency category changes with additional information.

Figure 3. Testing Strategy for Chlorobenzene



Potency Category	Probability
Nonsensitizer	0.22
Weak	0.27
Moderate	0.24
Strong/Extreme	0.27

Potency Category	Probability
Nonsensitizer	0.82
Weak	0.084
Moderate	0.072
Strong/Extreme	0.028

Potency Category	Probability
Nonsensitizer	0.92
Weak	0.049
Moderate	0.00097
Strong/Extreme	0.031

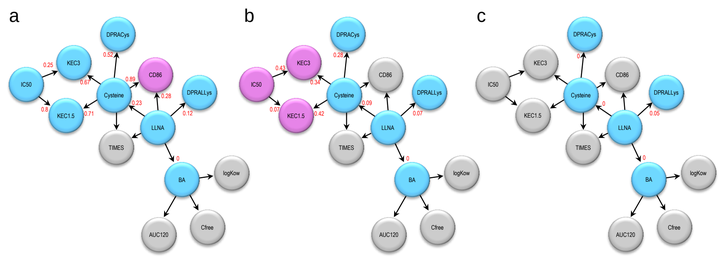
The abbreviations for the variables are listed in **Table 1**, except for BA = bioavailability. Blue indicates undefined variables, purple indicates the variables with the highest mutual information (shown in red type), and gray indicates variables with known values.
(a) With no information on chlorobenzene, the variable with the highest mutual information is TIMES.
(b) When the TIMES, logK_{ow}, and bioavailability (Cfree and AUC120) are known, the KeratinoSens data have the highest mutual information for the latent variable Cysteine.
(c) After KeratinoSens data are applied, the mutual information for the remaining variables is small.

- When the data for all of the variables are applied, the probability for the nonsensitizer category increases from that shown in **Figure 3c** by a small amount, to 0.97.

Case Study 2. 2-Mercaptobenzothiazole

- 2-Mercaptobenzothiazole is a moderate sensitizer (ICCVAM 2009).
- Testing Strategy
  - **Figure 4** shows how the mutual information values of the variables change as information is added. The accompanying tables show how the probability for each potency category changes with additional information.

Figure 4. Testing Strategy for 2-Mercaptobenzothiazole



Potency Category	Probability
Nonsensitizer	0.07
Weak	0.13
Moderate	0.43
Strong/Extreme	0.37

Potency Category	Probability
Nonsensitizer	0.011
Weak	0.069
Moderate	0.61
Strong/Extreme	0.31

Potency Category	Probability
Nonsensitizer	0.000045
Weak	0.036
Moderate	0.67
Strong/Extreme	0.29

The abbreviations for the variables are listed in **Table 1**, except for BA = bioavailability. Blue indicates undefined variables, purple indicates the variables with the highest mutual information (shown in red type), and gray indicates variables with known values.
(a) When the TIMES, logK_{ow}, and bioavailability (Cfree and AUC120) are known, the CD86 data have the highest mutual information for the LLNA. After the CD86 data are applied, the highest mutual information for the LLNA is yielded by the latent variable Cysteine.
(b) KeratinoSens data have the highest mutual information for Cysteine.
(c) After KeratinoSens data are added, the mutual information for the remaining variable with value for the LLNA, DPRALys, is small.

- When the data for all variables are included, the probability for the moderate category increases again slightly, compared with **Figure 4c**, to 0.71.

Acknowledgements

The Intramural Research Program of the National Institute of Environmental Health Sciences (NIEHS) supported this poster. Technical support was provided by ILS under NIEHS contracts N01-ES 35504 and HHSN27320140003C and SSS, Inc., under NIEHS contract GS-23F-9806H.

The views expressed above do not necessarily represent the official positions of any Federal agency. Since the poster was written as part of the official duties of the authors, it can be freely copied.

A summary of NICEATM and ICCVAM activities at the Ninth World Congress is available on the National Toxicology Program website at http://ntp.niehs.nih.gov/go/41583.